

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: October 3, 2005, 19:45:51 : Search time 1246 Seconds
(without alignments)
7120.774 Million cell updates/sec

Title: US-09-982-091A-2

Perfect score: 6604

Sequence: 1 MALCEESEGVFLPEPDISLK.....RFRKDSPTVKSRSIFQLLE 1285

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7442561 seqs, 3452328358 residues

Total number of hits satisfying chosen parameters: 14885122

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame.p2n.model -DEV=x1h
-CGN2_1/USPTO_EPO01/US09982091/runat_30092005_162543_10615/app_query.fasta_1.1479
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=0.0 -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pro NOR=ext -HEATSIZE=500 -MILEN=0
-MAXLEN=200000000 -USER=US09982091.0CGN_1.1.560.0/runat_30092005_162543_10615
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/PTU6_PUBCOMB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10J_PUBCOMB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
27: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6604	100.0	4754	9	US-09-982-091A-1
2	2695	40.8	4804	20	US-10-357-9920-30217
3	2687	40.7	4756	9	US-09-982-091A-3
4	784.5	11.9	5439	24	US-11-097-143-857
5	711.5	10.8	9315	24	US-11-097-143-856
6	498.5	7.5	111206	13	US-10-087-192-1957
7	452	6.8	58837	9	US-09-982-091A-5
8	399.5	6.0	20448	24	US-11-097-143-27224
9	393.5	6.0	547	9	US-09-998-598-480
10	384.5	5.8	484	9	US-09-998-598-268
11	382.5	5.8	9953	24	US-11-097-143-29297
12	376.5	5.7	10217	21	US-10-956-157-4592
13	373	5.6	7479	19	US-10-697-526-1
14	372	5.6	10330	21	US-10-956-157-426
15	372	5.6	10330	22	US-10-934-958-25
16	371	5.6	10452	11	US-09-968-007A-490
17	371	5.6	10452	21	US-10-843-641A-6960
18	371	5.6	10452	22	US-10-934-998-50
19	371	5.6	11167	16	US-10-252-157-188
20	366.5	5.5	3489	11	US-09-894-273-1
21	366.5	5.5	3489	15	US-10-294-804-1
22	366.5	5.5	3489	20	US-10-194-046-1
23	361.5	5.5	24789	24	US-11-097-143-40201
24	356	5.4	3705	21	US-10-741-849-6132
25	354.5	5.4	4380	17	US-10-094-466-29
26	354	5.4	12629	24	US-11-097-143-29296
27	349.5	5.3	18506	24	US-11-097-143-5629
28	348.5	5.3	4721	24	US-11-097-143-29150
29	346	5.2	5331	9	US-09-764-176-6
30	345.5	5.2	7097	24	US-11-097-143-29149
31	341	5.1	8910	24	US-11-097-143-2333
32	338.5	5.1	4935	19	US-10-437-963-47851
33	335.5	5.1	6665	20	US-10-335-053-229
34	334.5	5.1	6900	14	US-10-171-311-163
35	334.5	5.1	6900	22	US-10-764-445-13
36	333.5	5.0	7694	16	US-10-096-534-34
37	333.5	5.0	7695	11	US-09-968-007A-462
38	333.5	5.0	7695	21	US-10-843-641A-6932
39	333.5	5.0	8063	10	US-09-814-353-21776
40	333.5	5.0	24971	24	US-11-097-143-27223
41	331.5	5.0	6861	14	US-10-171-311-161
42	331.5	5.0	6861	17	US-10-341-434-102
43	331.5	5.0	6861	22	US-10-923-035-18
44	331.5	5.0	9373	24	US-11-097-143-7772
45	331.5	5.0	12505	24	US-11-097-143-7771

ALIGNMENTS

RESULT 1
US-09-982-091A-1
; Sequence 1, Application US/09982091A
; Patent No. US20020151030A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
; APPLICANT: KUMAGAI, AKIO
; APPLICANT: DUNPHY, WILLIAM
; TITLE OF INVENTION: CLASSIN PROTEINS AND METHODS OF USE THEREOF
; FILE REFERENCE: CIT11320-1
; CURRENT APPLICATION NUMBER: US/09/982,091A
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: US 60/241,246
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1

Pred. No. is the number of results predicted by chance to have a

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: October 3, 2005, 16:09:35 ; Search time 339 Seconds
(without alignments)
6202.406 Million cell updates/sec

Title: US-09-982-091a-2

Perfect score: 6604
Sequence: 1 MALCBBEQVLEPBDISLK.....RFRKSTPTVKSRISIFOLLE 1285

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-O=/cg2n2_1/USPTO_spool/US09982091/runat.30092005.162541.10527/app.query.fasta.1.1479
-DB=Issued_Patents_NA -QFMT=Istep -SUFFIX=p2n.rn1 -MINMATCH=0.1 -DOOPCL=50
-DOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORMe=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09982091 @CGN 1.1.68 @runat.30092005.162541.10527 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -IONBLOC
-DEV_TIMOUT=120 -WARN_TIMOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: Issued_Patents_NA.*
- 2: /cg2n2_6/prodata/1/ina/5A_COMB.seq.*
- 3: /cg2n2_6/prodata/1/ina/5B_COMB.seq.*
- 4: /cg2n2_6/prodata/1/ina/5A_COMB.seq.*
- 5: /cg2n2_6/prodata/1/ina/5B_COMB.seq.*
- 6: /cg2n2_6/prodata/1/ina/backfillseq1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2695.5	40.8	4756	4	US-09-949-016-4455
2	468	7.1	53737	4	US-09-949-016-16197
3	366.5	5.5	3489	2	US-08-728-323A-1
4	366.5	5.5	3489	3	US-09-298-568-1
5	366.5	5.5	3489	4	US-09-410-399-1
6	366.5	5.5	3489	4	US-09-894-273-1
7	366.5	5.5	32207	2	US-08-770-379-20
8	366.5	5.5	32207	3	US-08-757-669A-20
9	366.5	5.5	32207	3	US-09-230-371A-20
10	346	5.2	5331	4	US-09-764-176-6
11	345	5.2	6755	4	US-08-931-999-4
12	343.5	5.2	9551	1	US-08-056-200-93

13	343.5	5.2	9551	2	US-08-800-644-93	Sequence 93, Appl
14	331.5	5.0	6661	4	US-09-949-016-1240	Sequence 1240, Ap
15	331.5	5.0	6661	4	US-09-949-016-1241	Sequence 1241, Ap
16	331.5	5.0	6661	4	US-09-949-016-1242	Sequence 1242, Ap
17	321.5	4.9	8590	4	US-09-949-016-5562	Sequence 5562, Ap
18	321.5	4.9	10300	4	US-09-949-016-636	Sequence 636, App
19	317.5	4.8	38575	4	US-09-949-016-17304	Sequence 17304, A
20	317.5	4.8	119153	4	US-09-949-016-12378	Sequence 12378, A
21	308	4.7	6644	4	US-08-875-435B-5	Sequence 5, Appl1
22	303.5	4.6	5361	3	US-08-973-463-2	Sequence 2, Appl1
23	303.5	4.6	6152	3	US-08-978-463-1	Sequence 1, Appl1
24	300.5	4.6	5200	4	US-08-978-277A-3	Sequence 3, Appl1
25	299.5	4.5	9626	4	US-09-150-867-2	Sequence 2, Appl1
26	296.5	4.5	8503	4	US-09-620-312D-130	Sequence 130, App
27	295.5	4.5	6452	4	US-09-949-016-597	Sequence 597, App
28	295.5	4.5	6453	4	US-09-949-016-1533	Sequence 1533, Ap
29	295.5	4.5	8789	1	US-08-328-254-5	Sequence 5, Appl1
30	293.5	4.4	7453	4	US-09-620-312D-248	Sequence 248, App
31	293.5	4.4	7501	4	US-09-620-312D-249	Sequence 249, App
32	293	4.4	8257	4	US-09-595-684B-30	Sequence 30, Appl
33	292	4.4	5574	4	US-09-917-254-40	Sequence 40, Appl
34	292	4.4	5883	4	US-09-949-016-5001	Sequence 5001, Ap
35	292	4.4	6011	4	US-09-949-016-3017	Sequence 3017, Ap
36	290.5	4.4	7596	4	US-09-023-655-1463	Sequence 1463, Ap
37	290	4.4	7571	4	US-09-949-016-4366	Sequence 4366, Ap
38	290	4.4	8146	4	US-09-976-594-725	Sequence 725, Appl
39	290	4.4	11917	4	US-09-566-921-32	Sequence 32, Appl
40	290	4.4	101300	4	US-09-949-016-16108	Sequence 16108, A
41	280	4.4	145287	4	US-09-949-016-13530	Sequence 13530, A
42	280	4.4	145287	4	US-09-949-016-13531	Sequence 13531, A
43	286	4.3	6175	4	US-08-875-435B-1	Sequence 1, Appl1
44	284.5	4.3	10136	1	US-08-353-700-2	Sequence 2, Appl1
45	284.5	4.3	10136	5	PCT-US95-16216-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-09-949-016-4455
Sequence 4455 Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949, 016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4455
LENGTH: 4756
TYPE: DNA
ORGANISM: Human
US-09-949-016-4455

Alignment Scores:
Pred. No.: 4,42e-272
Score: 2695.50
Percent Similarity: 60.99%
Best Local Similarity: 46.14%
Query Match: 40.82%
DB: 4
Length: 4756
Matches: 640
Conservative: 206
Mismatch: 309
Indels: 232
Gaps: 44

US-09-982-091a-2 (1-1285) x US-09-949-016-4455 (1-4756)
Qy 9 GlnValPheLeuGlu-----ProGluApIleSerLeuYtIleValGlnThrAsp 25

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: October 3, 2005, 14:44:20 ; Search time 6050 Seconds

(without alignments)
8084.725 Million cell updates/sec

Title: US-09-982-091A-2

Perfect score: 6604
Sequence: 1 MALACBEEVFLPEPDISIK.....RFRKSTPTVKSRSIFOLLE 1285

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+g2n.model -DEV=x1h
-Q=/cgn2.1/USPTO/US09982091/runac_30092005.162541.10508/app_query.fasta.1.1479
-DB=EST -QFMT=fasta -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOEXT=0 -LOOEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR SCORE=Pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=Pct -NORM=EXT -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09982091@cgn2.1.1.3680@runac_30092005.162541.10508 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEX=7

Database :

EST: +
1: gb_ests1: +
2: gb_ests2: +
3: gb_hic: +
4: gb_ests3: +
5: gb_ests4: +
6: gb_ests5: +
7: gb_ests6: +
8: gb_g881: +
9: gb_g882: +

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1379	20.9	961	6	CA983046 AGENCOURT
2	1339	20.3	930	7	CA981092 AGENCOURT
3	1288.5	19.5	853	7	CR585850 CR585850
4	1189	18.0	922	6	CA985671 AGENCOURT
5	1105.5	16.7	933	6	BX756585 BX756585
6	1081	16.4	981	6	CA974658 AGENCOURT
7	1069	16.2	2075	3	BC018670 Homo sapi
8	1056	16.0	864	5	BX783588 BX783588
9	1030	15.6	707	1	AL859388 AL859388

10	1011	15.3	607	4	BJ041912 BJ041912
11	1003	15.2	915	6	CA791641 AGENCOURT
12	1002	15.2	2063	3	AK053847 AK053847
13	1001	15.2	634	1	AL886670 AL886670
14	998.5	15.1	926	5	BU097710 AGENCOURT
15	992	15.0	689	1	AL865075 AL865075
16	988	15.0	649	1	AL864962 AL864962
17	975.5	14.8	640	1	AL892911 AL892911
18	975	14.8	647	1	AL878964 AL878964
19	960.5	14.5	670	1	AL862969 AL862969
20	960	14.5	563	4	BJ029537 BJ029537
21	922	14.0	548	2	AW67107 AW67107
22	905	13.7	858	1	AL863727 AL863727
23	893.5	13.5	955	6	CA982465 CA982465
24	874	13.2	527	2	BF427206 BF427206
25	850	12.9	574	1	AL857623 AL857623
26	830.5	12.6	751	7	CF749459 CF749459
27	829.5	12.6	929	5	BO963412 BO963412
28	804	12.2	864	7	CR416889 CR416889
29	790.5	12.0	1027	5	BQ072224 BQ072224
30	786	11.9	471	2	BE189872 BE189872
31	735.5	11.1	663	1	AL891701 AL891701
32	728.5	11.0	700	7	CK635361 CK635361
33	725	11.0	532	5	BX716882 BX716882
34	719	10.9	537	4	BI449197 BI449197
35	712	10.8	420	2	BF612124 BF612124
36	711	10.8	649	1	AL863524 AL863524
37	708.5	10.7	643	1	AL957528 AL957528
38	706	10.7	631	1	AL850776 AL850776
39	698.5	10.6	630	7	CF724962 CF724962
40	691.5	10.5	691	7	CF748227 CF748227
41	668.5	10.1	661	1	AL870060 AL870060
42	665	10.1	624	1	AJ392625 AJ392625
43	657	9.9	672	1	AJ394263 AJ394263
44	655	9.9	590	2	AW645447 AW645447
45	646	9.8	582	5	BP313518 BP313518

ALIGNMENTS

RESULT 1
CA983046
LOCUS
DEFINITION CA983046 961 bp mRNA linear EST 27-FEB-2003
IMAGE:6870256 5', mRNA sequence.

ACCESSION
CA983046
VERSION
CA983046.1 GI:27515700

KEYWORDS
SOURCE
ORGANISM
Xenopus laevis (African clawed frog)

EST.
Xenopus laevis
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus; Xenopus.
1 (bases 1 to 961)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: N. Garret, P. Lemaire, A.M. Zorn, and J.B.
Gurdon (Wellcome/CRC Institute)
CDNA library Preparation: N. Garret, P. Lemaire, A.M. Zorn, and
J.B. Gurdon (Wellcome/CRC Institute)
CDNA library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: XSC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
Plate: LLM14500 row: f column: 15
High quality sequence start: 11
High quality sequence stop: 710.
Location/Qualifiers

FEATURES

AA 17-OCT-2000; 2000US-0241246P
PR

